Instituto Complutense de Análisis Económico

UNIVERSIDAD COMPLUTENSE

FACULTAD DE ECONOMICAS

Campus de Somosaguas

28223 MADRID

Teléfono 394 26 11 - FAX 294 26 13

Documento de trabajo

A Generalizad Least Squares Estimation Method for Varma Models

Rafael Flores de Frutos

Gregorio R. Serrano

No.9712 Junio 1997

Instituto Complutense de Análisis Económico

UNIVERSIDAD COMPLUTENSE

A GENERALIZED LEAST SQUARES ESTIMATION METHOD FOR VARMA MODELS

Rafael Flores de Frutos
Gregorio R. Serrano

Departamento de Economía Cuantitativa e ICAE
Universidad Complutense de Madrid
Campus de Somosaguas, 28223 Madrid

ABSTRACT

In this paper a new generalized least squares procedure for estimating VARMA models is proposed. This method differs from existing ones in explicitly considering the stochastic structure of the approximation error that arises when lagged innovations are replaced with lagged residuals obtained from a long VAR. Simulation results indicate that this method improves the accuracy of estimates with small and moderate sample sizes, and increases the frequency of identifying small nonzero parameters, with respect to both Double Regression and exact maximum likelihood estimation procedures.

RESUMEN

En este artículo se propone un nuevo método lineal para la estimación de modelos VARMA. Este método se diferencia de otros en considerar explícitamente el error que se comete al aproximar las innovaciones a través de los residuos minimocuadráticos procedentes de un VAR largo. Los resultados de un ejercicio de simulación revelan que el método mejora la precisión de las estimaciones, en muestras pequeñas y moderadas, con respecto al método de Doble Regresión y máxima verosimilitud exacta. También aumenta la frecuencia con que se detectan parámeros pequeños en tareas de identificación.

Keywords. VARMA models estimation; generalized least squares; model specification

1. INTRODUCTION

While it is recognized that in some situations a mixed VARMA model might produce better forecasts than an appropriate finite order VAR approximation, the fact is that VAR models have dominated the empirical work. The painstaking specification and estimation procedures associated to VARMA models, along with the lack of evidence about their superior forecasting performance, help to understand the choice made by many econometricians.

Simplifying the task of elaborating VARMA models has been the goal of many authors. Some of them, as Spliid (1983), Hanan and Kavalieris (1984), Koreisha and Pukkila (1989) or Reinsel et al. (1992), have developed linear estimation procedures with some desirable features:

i) They are easy to implement; most of then only require a standard least squares (LS) routine.

ii) They are fast; either no iterations or just a few are needed for obtaining accurate estimates, comparable with that of maximum likelihood (ML) methods.

iii) For the univariate case, Koreisha and Pukkila (1990) have found that their generalized least squares (GLS) procedure: (1) yields accurate estimations even when short samples are used, (2) seldom generates non-invertible or non-stationary situations, and (3) performs better than ML when a pure moving average (MA) process is fitted to a short sample.

iv) They are useful in identification tasks; fast estimation procedures have proved to be quite effective in detecting nonzero parameters as well as in finding the orders p and q of the AR and MA polynomial matrices.

v) Using these estimates to initialize exact ML procedures helps to reduce the number of iterations as well as produces more reliable final estimates.

In this line, we propose a new GLS based method for estimating VARMA models. We use the idea, introduced by Koreisha and Pukkila (1990) in the univariate context, of taking into account the approximation error from replacing, in the original VARMA model, lagged innovations with lagged residuals obtained from a long VAR(L). Instead of using the Koreisha and Pukkila's white noise assumption about the approximation error, we derive its stochastic structure, and show that it depends on "L", the order of the long VAR, as well as on the orders "p" and "q" of the VARMA model. This structure induces a VARMA process in the noise of the model to be estimated.

For the univariate case, our procedure is asymptotically equivalent to that proposed by Koreisha and Pukkila (1990). For the multivariate case, it is asymptotically equivalent to the procedures proposed in Koreisha and Pukkila (1989). However, we show that in specification tasks our estimator may increase the power of the standard t-test in detecting nonzero parameters. When compared with the standard Double Regression method, simulation results indicate that our method yields more accurate estimates and shows a better performance in detecting small nonzero parameters. The same simulation results show how our method may yield as accurate estimates as those from exact maximum likelyhood procedures.

The paper is organized as follows. Section 2 first describes our proposed GLS approach for estimating VMA processes, then this procedure is extended to general VARMA models. Section 3 presents a simulation exercise. Finally, Section 4 summarizes the main conclusions.

2. A NEW GLS APPROACH FOR ESTIMATING VARMA MODELS

2.1 The case of pure VMA models

Consider a $k\times 1$ vector $\mathbf{z}_t$ of time series following the invertible VMA process:

$$\mathbf{z}_t = \theta_q(\mathbf{B})\ \mathbf{a}_t \tag{1}$$

$t= 1, 2, \ldots N$, where $\theta_q(\mathbf{B}) = \mathbf{I} - \theta_1\mathbf{B} - \ldots - \theta_q\mathbf{B}^q$ is a $k\times k$ finite order (q) polynomial matrix in the lag operator $\mathbf{B}$, with the roots of $|\ \theta_q(\mathbf{B})| = \mathbf{0}$ lying outside the unit circle. The $k\times 1$ vector $\mathbf{a}_t$ is assumed to follow a white noise process with covariance matrix Σ_a.

The infinite VAR representation of (1) is:

$$\mathbf{z}_t = \sum_{j=1}^{\infty} \pi_j\ \mathbf{z}_{t-j} + \mathbf{a}_t \tag{2}$$

Due to the invertibility of (1) π_j approaches to zero as j approaches to infinite, therefore a long but finite VAR(L) process might be a good approximation for (1):

$$\mathbf{z}_t = \sum_{j=1}^{L} \pi_j\ \mathbf{z}_{t-j} + \mathbf{u}_t \tag{3}$$

The choice of L, the order of the VAR in (3), should be based on the data at hand. For nonseasonal data, a value of L between log N and $\sqrt{N}$ is considered reasonable for many authors. See Lütkepohl and Poskitt (1996), page 73, for a recent discussion on this topic.

From (2) and (3) is easy to obtain that:

$$\begin{aligned} a_t &= u_t + \epsilon_t \\ \epsilon_t &= S_{1t} - S_{2t} \end{aligned} \tag{4}$$

where

$$\begin{aligned} S_{1t} &= \sum_{i=1}^{L} \pi_i \, z_{t-i} \\ S_{2t} &= \sum_{i=1}^{\infty} \pi_i \, z_{t-i} \end{aligned} \tag{5}$$

S_{1t} and S_{2t} can also be expressed as:

$$\begin{aligned} S_{1t} &= [\, I - \pi_L(B) \,] \, z_t \\ S_{2t} &= z_t - a_t \end{aligned} \tag{6}$$

where

$$\pi_L(B) = [\, I - \pi_1 B - \dots - \pi_L B^L \,] \tag{7}$$

Using (6) the approximation error ϵ_t can be expressed as:

$$\epsilon_t = a_t - \pi_L(B) \, z_t = [\, I - \pi_L(B) \, \theta_q(B) \,] \, a_t \tag{8}$$

By substituting (8) into (4), and (4) into (1) we obtain:

$$z_t = [\, \theta_q(B) - I \,] \, u_t + \eta_t \tag{9}$$

where η_t follows the VMA(2q+L) process:

$$\begin{aligned} \eta_t &= [\, \theta_q(B) - (\theta_q(B) - I) \, \pi_L(B) \, \theta_q(B) \,] \, a_t \\ &= [\, I - \Psi_1 B - \dots - \Psi_{2q+L} B^{2q+L} \,] \, a_t \\ &= \Psi_{2q+L}(B) \, a_t \end{aligned} \tag{10}$$

For all observations, model (9)-(10) can be expressed as:

$$\begin{aligned} \mathrm{vec}(Z_N) &= (U' \otimes I_k) \, \mathrm{vec}(\theta) + \mathrm{vec}(H_N) \\ \mathrm{vec}(H_N) &= D_{\Psi,N} \, \mathrm{vec}(A_N) + G_{\Psi,N} \, \mathrm{vec}(A^{**}) \end{aligned} \tag{11}$$

with

$$Z_N = [z_1 \; z_2 \; \dots \; z_N]_{(k \times N)} \tag{12}$$

$$U = \begin{bmatrix} u_0 & u_1 & \dots & u_{N-1} \\ u_{-1} & u_0 & \dots & u_{N-2} \\ .. & .. & \dots & .. \\ u_{-q+1} & u_{-q+2} & \dots & u_{N-q} \end{bmatrix}_{(qk \times N)} \tag{13}$$

$$\theta = [-\theta_1 \; -\theta_2 \; \dots \; -\theta_q]_{(k \times qk)} \tag{14}$$

$$\begin{aligned} H_N &= [\, \eta_1 \; \eta_2 \; \dots \; \eta_N \,]_{(k \times N)} \\ A^{**} &= [a_{-2q+L+1} \; \dots \; a_0 \,]_{(k \times 2q+L)} \end{aligned} \tag{15}$$

$$A_N = [a_1 \; a_2 \; \dots \; a_N]_{(k \times N)} \tag{16}$$

$$D_{\psi,N} = \begin{bmatrix} I_k & 0 & 0 & .. & 0 & 0 & .. & 0 & 0 & 0 \\ -\psi_1 & I_k & 0 & .. & 0 & 0 & .. & 0 & 0 & 0 \\ .. & -\psi_1 & I_k & .. & 0 & 0 & .. & 0 & 0 & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ -\psi_{2q+L} & .. & .. & -\psi_1 & I_k & 0 & .. & 0 & 0 & 0 \\ 0 & -\psi_{2q+L} & .. & .. & -\psi_1 & I_k & .. & 0 & 0 & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & 0 & 0 & .. & .. & .. & -\psi_1 & I_k & 0 & 0 \\ 0 & 0 & 0 & .. & -\psi_{2q+L} & .. & .. & -\psi_1 & I_k & 0 \\ 0 & 0 & 0 & .. & 0 & -\psi_{2q+L} & .. & .. & -\psi_1 & I_k \end{bmatrix}_{(Nk \times Nk)} \tag{17}$$

$$
\mathbf{G}_{\psi,N} = \begin{bmatrix} -\psi_{2q+L} & .. & .. \; .. & -\psi_2 & -\psi_1 \\ 0 & -\psi_{2q+L} & .. \; .. & .. & -\psi_2 \\ .. & .. & .. \; .. & .. & .. \\ 0 & .. & .. \; .. & -\psi_{2q+L} & .. \\ 0 & 0 & .. \; .. & 0 & -\psi_{2q+L} \\ 0 & 0 & .. \; .. & .. & 0 \\ .. & .. & .. \; .. & .. & .. \\ 0 & 0 & .. \; .. & 0 & .. \\ 0 & 0 & .. \; .. & 0 & 0 \end{bmatrix}_{(kN\times(2q+L)k)} \tag{18}
$$

As $\mathbf{vec(H_N)}$ has not a scalar covariance matrix, a feasible GLS approach seems to be adequate for the estimation of **(11)**:

$$
\mathbf{vec}(\hat{\theta}) = [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecH_N}(\hat{U}'\otimes I_k)]^{-1} (\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecH_N}\mathbf{vec}(Z_N) \tag{19}
$$

whose asymptotic variance-covariance matrix can be estimated as:

$$
\hat{\Sigma}_{GHR} = [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecH_N}(\hat{U}'\otimes I_k)]^{-1} \tag{20}
$$

where

$$
\hat{\Sigma}_{vecH_N} = (\hat{G}_{\Psi,N} \,\vdots\, \hat{D}_{\Psi,N}) \, (I_{N+2q+L} \otimes \hat{\Sigma}_a) \, (\hat{G}_{\Psi,N} \,\vdots\, \hat{D}_{\Psi,N})' \tag{21}
$$

The estimation of $\mathbf{U}$ can be obtained using the residuals from a LS fit to **(3)**. A consistent estimation of $\mathbf{vec}(\theta)$, necessary for estimating $\mathbf{D}_{\psi,N}$ and $\mathbf{G}_{\psi,N}$, can be obtained by applying LS to **(11)** once $\mathbf{U}$ has been replaced with $\hat{\mathbf{U}}$. Finally, an estimation of Σ_a can be obtained as:

$$
\hat{\Sigma}_a = \frac{\hat{U}_N\hat{U}_N'}{N} \tag{22}
$$

or from:

$$
\mathbf{vec}(\hat{\Sigma}_a) = (I + \hat{\Psi}_1 \otimes \hat{\Psi}_1' + \ldots + \hat{\Psi}_{2q+L} \otimes \hat{\Psi}_{2q+L}')^{-1} \, \mathbf{vec}(\hat{\Sigma}_{H_N}) \tag{23}
$$

Note that making **L** to depend on **N**, it allows to prove that: **(i)** LS to **(3)** yields consistent estimates of π_j, **(ii)** the residuals series $\hat{u}_t$ approach in probability the true innovation series a_t, and therefore, **(iii)** LS to **(11)** yields consistent estimates of $\mathbf{vec}(\theta)$. See Koreisha and Pukkila (1989), page 329 or Lütkepohl (1993), pages 268 and 306.

Our proposed estimation procedure for model **(11)** can be summarized as follows:

1) Get initial estimates for $\mathbf{u}_t$, Σ_a and $\pi_L(\mathbf{B})$ by applying LS to:

$$
z_t = \sum_{j=1}^{L} \pi_j \, z_{t-j} + u_t \tag{24}
$$

2) Compute $\hat{\mathbf{U}}$ from $\hat{\mathbf{u}}_t$ and get an estimation of $\mathbf{vec}(\theta)$ by applying LS to:

$$
\mathbf{vec}(Z_N) = [\hat{U}' \otimes I_k] \, \mathbf{vec}(\theta) + \mathbf{vec}(H_N) \tag{25}
$$

3) Estimate $\mathbf{G}_{\psi,N}$ and $\mathbf{D}_{\psi,N}$ using the coefficients in:

$$
\hat{\Psi}_{2q+L}(B) = [\, \hat{\theta}_q(B) - (\hat{\theta}_q(B) - I) \, \hat{\pi}_L(B) \, \hat{\theta}_q(B) \,] \tag{26}
$$

4) Find the Choleski factor matrix $\mathbf{T}$ for:

$$
\hat{\Sigma}_{vecH_N} = (\hat{G}_{\Psi,N} \,\vdots\, \hat{D}_{\Psi,N}) \, (I_{N+2q+L} \otimes \hat{\Sigma}_a) \, (\hat{G}_{\Psi,N} \,\vdots\, \hat{D}_{\Psi,N})' \tag{27}
$$

5) Apply LS to:

$$
T^{-1} \, \mathbf{vec}(Z_N) = T^{-1} \, (\hat{U}' \otimes I_k) \, \mathbf{vec}(\theta) + T^{-1} \, \mathbf{vec}(H_N) \tag{28}
$$

It is important to mention that:

1) In Koreisha and Pukkila(1989) the approximation error ϵ_t is implicitly assumed to be 0 and LS to **(11)** is the proposed estimation procedure, this is called the Double Regression (DR) estimation method. The expressions for the DR estimator and its variance-covariance matrix are:

$$
\mathbf{vec}(\hat{\theta}) = [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecA_N}(\hat{U}'\otimes I_k)]^{-1} (\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecA_N}\mathbf{vec}(Z_N) \tag{29}
$$

$$
\hat{\Sigma}_{DR} = [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecA_N}(\hat{U}'\otimes I_k)]^{-1} \tag{30}
$$

where

$$
\hat{\Sigma}_{vecA_N} = (I_N \otimes \hat{\Sigma}_a) \tag{31}
$$

Note that instead of using the usual expression:

$$
\hat{\Sigma}_{LS} = (\hat{U}\hat{U}'\otimes I_k)^{-1}(\hat{U}\otimes I_k)\hat{\Sigma}_{vecA_N}(\hat{U}'\otimes I_k)(\hat{U}\hat{U}'\otimes I_k)^{-1} \tag{32}
$$

for the variance-covariance matrix of the LS estimator, DR uses **(30)** that is a consistent estimator (by the method of moments) of the asymptotic variance-covariance matrix of **(29)**.

2) Koreisha and Pukkila(1990), into a univariate framework, assume a white noise process for ϵ_t and propose a GLS approach based on a MA(q) process for $\mathbf{vec(H_N)}$. The extension of this method to the multivariate framework relies on applying GLS to **(25)** where $\mathbf{vec(H_N)}$ is assumed to follow a VMA(q) process.

We have shown that ϵ_t does not follow a white noise process but a VMA(L+q-1), what implies that $\mathbf{vec(H_N)}$ follows a VMA(2q+L). As Koreisha and Pukkila(1990) we propose a GLS estimation procedure, but taking into account the exact structure for $\mathbf{vec(H_N)}$. We call this the Generalized Hannan Rissanen (GHR) estimation procedure.

3) DR and GHR estimators have the same asymptotic distribution, note that $\mathbf{plim}\ \hat{\mathbf{u}}_t = \mathbf{a}_t$, $\mathbf{plim}\ \mathbf{G} = \mathbf{0}$ and $\mathbf{plim}\ \mathbf{D} = \mathbf{I}_{Nk}$, which can be obtained from:

$$\sqrt{n}\ [\mathrm{vec}(\hat{\Theta}) - \mathrm{vec}(\Theta)] \rightarrow \mathrm{N}\ [0\ ,\ (\mathrm{I}_q \otimes \Sigma_a)^{-1} \otimes \Sigma_a] \tag{33}$$

However, the expression **(20)**, i.e. the estimator of the asympthotic variance-covariance matrix for the vector of parameters estimator, has some desirable features:

(a) The difference **(32)-(20)** is positive semidefinite, see Judge et al. (1982) pp. 293-294 for the standard proof. Thus, in finite samples **(20)** will produce in general lower standard errors than **(32)**. If **(20)**, instead of **(32)**, is used for testing the significance of a particular parameter, the power of the standard t-test increases.

(b) The difference **(30)-(20)**, will be positive semidefinite if:

$$\hat{\Sigma}_{\mathrm{vecA}_N} - \hat{\Sigma}_{\mathrm{vecH}_N} \tag{34}$$

is a positive semidefinite matrix too[1]. In such cases, the power of the standard t-test can be augmented by using the GHR expression **(20)**. The result in (b) is important because it shows that **(20)** might be preferable to **(30)**. In specification tasks, both expressions **(20)** and **(30)** can be used together in order to minimize the probability of removing significant parameters.

2.2 The case of mixed VARMA models

Consider the VARMA(p,q) process:

$$\phi_p(\mathrm{B})\ \mathrm{z}_t = \theta_q(\mathrm{B})\ \mathrm{a}_t \tag{35}$$

$t = 1, 2, \dots$ N. Where $\phi_p(\mathbf{B}) = \mathbf{I} - \phi_1\mathbf{B} - \dots - \phi_p\mathbf{B}^p$ is a k×k finite order (**p**) polynomial matrix in the lag operator B, and with the roots of $|\phi(\mathbf{B})| = \mathbf{0}$ lying outside the unit circle. The remaining terms in **(35)** have been defined at the beginning of the previous section.

Consider the following two alternative representations of **(35)**:

<u>Representation #1</u>:

$$\begin{aligned} \mathbf{D}_{\phi,N}\ \mathrm{vec}(\mathbf{Z}_N) &= \mathbf{D}_{\theta,N}\ \mathrm{vec}(\mathbf{A}_N) + \mathbf{J}_{N,q^*}\ \nu \\ \nu &= \mathbf{C}_{\phi,q^*}\ \mathrm{vec}(\mathbf{Z}^*) - \mathbf{C}_{\theta,q^*}\ \mathrm{vec}(\mathbf{A}^*) \end{aligned} \tag{36}$$

<u>Representation #2</u>:

$$\mathrm{vec}(\mathbf{Z}_N) = (\mathbf{X}_a \otimes \mathbf{I}_k)\ \mathrm{vec}(\beta) + \mathrm{vec}(\mathbf{A}_N) \tag{37}$$

Where matrices $\mathbf{D}_{\phi,N}$ and $\mathbf{D}_{\theta,N}$ has the same structure as $\mathbf{D}_{\psi,N}$ but with elements $\phi_1, \dots \phi_q$, and $\theta_1, \dots \theta_q$, instead of $\Psi_1 \dots \Psi_{2q+L}$. The remaining matrices take the form:

$$\mathbf{J}_{N,q^*} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & .. & .. & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & .. & .. & \mathbf{0} & \mathbf{0} \\ .. & .. & .. & .. & .. & .. \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & \mathbf{0} \\ .. & .. & .. & .. & .. & .. \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & \mathbf{0} \end{bmatrix}_{(kN \times q^* k)} \tag{38}$$

$$\mathbf{C}_{\phi,q^*} = \begin{bmatrix} \phi_{q^*} & \phi_{q^*-1} & .. & .. & \phi_2 & \phi_1 \\ \mathbf{0} & \phi_{q^*} & .. & .. & \phi_3 & \phi_2 \\ .. & .. & .. & .. & .. & .. \\ \mathbf{0} & \mathbf{0} & .. & .. & \phi_{q^*} & \phi_{q^*-1} \\ \mathbf{0} & \mathbf{0} & .. & .. & \mathbf{0} & \phi_{q^*} \end{bmatrix}_{(q^* k \times q^* k)} \tag{39}$$

$$\mathbf{Z}^* = [\ \mathbf{z}_{-q^*+1}\ \mathbf{z}_{-q^*+2} \dots \mathbf{z}_{-1}\ \mathbf{z}_0\ ]_{(k \times q^*)} \tag{40}$$

$$\mathbf{A}^* = [\ \mathbf{a}_{-q^*+1}\ \mathbf{a}_{-q^*+2} \dots \mathbf{a}_{-1}\ \mathbf{a}_0\ ]_{(k \times q^*)} \tag{41}$$

$$\beta = [\ \phi_1\ \phi_2 \dots \phi_p\ -\theta_1\ -\theta_2 \dots -\theta_q\ ] \tag{42}$$

$$
\mathbf{X}_a = \begin{bmatrix} z_0 & z_1 & \cdots & z_{N-1} \\ z_{-1} & z_0 & \cdots & z_{N-2} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ z_{-p+1} & z_{-p+2} & \cdots & z_{N-p} \\ a_0 & a_1 & \cdots & a_{N-1} \\ a_{-1} & a_0 & \cdots & a_{N-2} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ a_{-q+1} & a_{-q+2} & \cdots & a_{N-q} \end{bmatrix}_{((p+q)k \times N)} \tag{43}
$$

Matrix $\mathbf{C}_{\theta,q^*}$ has the same structure that $\mathbf{C}_{\phi,q^*}$ but with elements $\theta_1, \dots \theta_{q^*}$ instead of $\phi_1, \dots \phi_{q^*}$. The order q^* is the maximum between p and q. Thus, if $q^*=q$, matrices ϕ_j will be zero for $j>p$. If $q^*=p$, matrices θ_j will be zero for $j>q$.

Now

$$
\Delta(\mathbf{B})\ \epsilon_t = \theta^{\epsilon}(\mathbf{B})\ \mathbf{a}_t \tag{44}
$$

where

$$
\begin{aligned} \Delta(\mathbf{B}) &= |\phi(\mathbf{B})|\ \mathbf{I}_k \\ &= \mathbf{I} - \Delta_1 \mathbf{B} - \Delta_2 \mathbf{B}^2 - \dots - \Delta_{p_\epsilon} \mathbf{B}^{p_\epsilon} \end{aligned} \tag{45}
$$

$$
\begin{aligned} \theta^{\epsilon}(\mathbf{B}) &= [\ \Delta(\mathbf{B}) - \pi_L(\mathbf{B})\ \phi^a(\mathbf{B})\ \theta_q(\mathbf{B})\ ] \\ &= -\theta^{\epsilon}_1 \mathbf{B} - \dots - \theta^{\epsilon}_{q_\epsilon} \mathbf{B}^{q_\epsilon} \end{aligned} \tag{46}
$$

and

$$
\begin{aligned} \phi^a(\mathbf{B}) &= \Delta(\mathbf{B})\ \phi^{-1}(\mathbf{B}) \\ &= \mathbf{I} - \phi^a_1 \mathbf{B} - \dots - \phi^a_{p(k-1)} \mathbf{B}^{p(k-1)} \end{aligned} \tag{47}
$$

That is, the approximation error ϵ_t follows a VARMA(p_ϵ, q_ϵ), where $p_\epsilon=pk$ and $q_\epsilon=L+q+p(k-1)$ will be the orders of $\Delta(\mathbf{B})$ and $\theta^{\epsilon}(\mathbf{B})$ respectively.

Using **(4)** and **(44)**, **(35)** can be expressed as:

$$
\begin{aligned} \phi_p(\mathbf{B})\ z_t &= [\ \theta_q(\mathbf{B}) - \mathbf{I}\ ]\ \mathbf{u}_t + \eta_t \\ \Delta(\mathbf{B})\ \eta_t &= \theta^{\eta}(\mathbf{B})\ \mathbf{a}_t \end{aligned} \tag{48}
$$

where

$$
\begin{aligned} \theta^{\eta}(\mathbf{B}) &= [\Delta(\mathbf{B}) + (\ \theta_q(\mathbf{B}) - \mathbf{I}\ )\ \theta^{\epsilon}(\mathbf{B})\ ] \\ &= \mathbf{I} - \theta^{\eta}_1 \mathbf{B} - \dots - \theta^{\eta}_{q_\eta} \mathbf{B}^{q_\eta} \end{aligned} \tag{49}
$$

q_η is $\max\{p_\epsilon,\ q+q_\epsilon\}$.

The whole sample representation of **(48)** is:

$$
\begin{aligned} \mathrm{vec}(\mathbf{Z}_N) &= (\mathbf{X}_u' \otimes \mathbf{I}_k)\ \mathrm{vec}(\beta) + \mathrm{vec}(\mathbf{H}_N) \\ \mathbf{D}_{\Delta,N}\ \mathrm{vec}(\mathbf{H}_N) &= \mathbf{D}_{\theta^\eta,N}\ \mathrm{vec}(\mathbf{A}_N) + \mathbf{J}_{N,q_\eta}\ \nu \\ \nu &= \mathbf{C}_{\Delta,q_\eta}\ \mathrm{vec}(\mathbf{H}^*) - \mathbf{C}_{\theta^\eta,q_\eta}\ \mathrm{vec}(\mathbf{A}^*) \end{aligned} \tag{50}
$$

where

$$
\mathbf{X}_u = \begin{bmatrix} z_0 & z_1 & \cdots & z_{N-1} \\ z_{-1} & z_0 & \cdots & z_{N-2} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ z_{-p+1} & z_{-p+2} & \cdots & z_{N-p} \\ u_0 & u_1 & \cdots & u_{N-1} \\ u_{-1} & u_0 & \cdots & u_{N-2} \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ u_{-q+1} & u_{-q+2} & \cdots & u_{N-q} \end{bmatrix}_{((p+q)k \times N)} \tag{51}
$$

$$
\mathbf{H}^* = [\ \eta_{-q_\eta+1}\ \eta_{-q_\eta+2}\ \cdots\ \eta_{-1}\ \eta_0\ ]_{(k \times q_\eta)} \tag{52}
$$

$$
\mathbf{A}^* = [\ \mathbf{a}_{-q_\eta+1}\ \mathbf{a}_{-q_\eta+2}\ \cdots\ \mathbf{a}_{-1}\ \mathbf{a}_0\ ]_{(k \times q_\eta)} \tag{53}
$$

and matrices $\mathbf{D}_{\Delta,N}$, $\mathbf{D}_{\theta\eta,N}$, $\mathbf{J}_{N,q\eta}$, $\mathbf{C}_{\Delta,q\eta}$ and $\mathbf{C}_{\theta\eta,q\eta}$ have the same structure as the corresponding $\mathbf{D}_{\phi,N}$, $\mathbf{D}_{\theta,N}$, $\mathbf{J}_{N,q^*}$, $\mathbf{C}_{\phi,q^*}$ and $\mathbf{C}_{\theta,q^*}$ but with elements and orders determined by polynomials $\Delta(\mathbf{B})$ and $\theta_q(\mathbf{B})$.

The structure of the model **(50)** suggests a GLS approach for estimating $\mathrm{vec}(\beta)$:

$$\mathrm{vec}(\hat{\beta}) = [(\hat{X}_u \otimes I_k)\hat{\Sigma}^{-1}_{\mathrm{vecH}_N}(\hat{X}_u' \otimes I_k)]^{-1} (\hat{X}_u \otimes I_k)\hat{\Sigma}^{-1}_{\mathrm{vecH}_N}\mathrm{vec}(Z_N) \tag{54}$$

whose variance-covariance matrix can be estimated as:

$$\hat{\Sigma}_{GHR} = [(\hat{X}_u \otimes I_k)\hat{\Sigma}^{-1}_{\mathrm{vecH}_N}(\hat{X}_u' \otimes I_k)]^{-1} \tag{55}$$

where

$$\begin{aligned}\hat{\Sigma}_{\mathrm{vecH}_N} = &(\hat{D}^{-1}_{\hat{\Delta},N} \times \hat{D}_{\theta',N}) (I \otimes \hat{\Sigma}_a) (\hat{D}^{-1}_{\hat{\Delta},N} \times \hat{D}_{\theta',N})' + \\ &+ (\hat{D}^{-1}_{\hat{\Delta},N} \times J_{N,q_*}) \Gamma (\hat{D}^{-1}_{\hat{\Delta},N} \times J_{N,q_*})'\end{aligned} \tag{56}$$

Matrix Γ represents the estimated variance-covariance matrix of the initial conditions "ν". Note that the last term in **(56)** disappears if "ν" are assumed to be 0.

Finally it is important to mention that matrices δ_k can be computed from:

$$\begin{aligned}\Delta(B) &= \prod_{i=1}^{pk} (I_k - \lambda_i I_k B) \\ &= I - \Delta_1 B - \dots - \Delta_{p_*} B^{p_*}\end{aligned} \tag{57}$$

where λ_i, $i = 1, 2, \dots pk$, are the eigenvalues of:

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}_{(pk \times pk)} \tag{58}$$

Also, matrices ϕ^a_j in $\phi^a(B)$ can be computed as:

$$\phi^a_j = \begin{cases} \Delta_j + \Delta_{j-1}\Psi_1 + \Delta_{j-2}\Psi_2 + \dots + \Delta_1\Psi_{j-1} - \Psi_j & \forall\, j \le (k-1)p \\ 0 & \forall\, j > (k-1)p \end{cases} \tag{59}$$

$$\Delta_s = 0 \quad \forall\, s \le 0$$

where $\Psi(B) = I + \Psi_1 B + \Psi_2 B^2 + \dots = \phi^{-1}(B)$.

Matrices Ψ_j can be computed as:

$$\begin{aligned}\Psi_j &= \phi_j + \phi_{j-1}\Psi_1 + \phi_{j-2}\Psi_2 + \dots + \phi_1\Psi_{j-1} \\ \phi_s &= 0 \quad \forall\, s > p \ \text{ or } \ s \le 0\end{aligned} \tag{60}$$

3. SIMULATION EXERCISE

Tables I and II and III show the simulation results for a VMA(1), a VMA(2) and a VARMA(1,1) process, respectively. These models are the same used in Koreisha and Pukkila (1989) for illustrating the properties of the Double Regression estimation procedure. We simulated 100 realizations for each model[2]. The sample size N was set equal to 100 and the order "L" of the long VAR was set equal to $\sqrt{N}=10$. Koreisha and Pukkila (1989) argue that the chosen models typify most practical real data applications, for instance: the density of nonzero elements is low, the variation in the magnitude of parameters values is broad and the feedback/causal mechanisms are complex.

All tables have the same structure, the first panel shows the mean value of parameter estimates obtained with three different estimation procedures: The Generalized Hannan Rissanen (GHR) procedure proposed in this paper, the Double Regression (DR) procedure proposed in Koreisha and Pukkila (1989) and the Exact Maximum Likelihood (EML) estimation procedure as proposed in Mauricio (1995)[3]. The second panel shows the mean values of the estimated standard errors associated to each parameter. The third panel shows the Mean Square Errors (MSE) computed from the two previous panels. Finally, the fourth panel shows the frequency of significant nonzero parameters (95% confidence) tentatively identified by each method.

Comparative results are:

1) All estimation procedures yield very similar parameters estimates. Only in the case of the VARMA(1,1) linear methods seem to underestimate some of the coefficients.

2) For VMA models, the estimated standard errors (s.e.) associated to DR estimates are always bigger than those associated to either GHR or EML. The later perform very similar. For the VARMA model, EML shows the biggest s.e.; GHR and DR yield similar results.

3) In terms of MSE and for VMA models, DR shows the lowest precision. Again, GHR and EML methods performs similarly. For the VARMA model, the EML method shows the highest MSE while the performance of DR and

GHR is similar. Given the results in 1) and 2) the problem seems to arise in the approximation used for the Information matrix.

4) If a standard t-static (with the usual 95% probability level) on the estimates in the first two panels, is used for identifying nonzero parameters, the GHR method is able to detect all relevant parameters in the VMA processes. Both DR and EML fail to detect the parameter .2 in the VMA(1) and the parameter .3 in the VMA(2). In the case of the VARMA model, none method is able to detect the presence of parameters .6 and -.5,; EML neither detects the presence of .4 and -1.1.

Now looking at the fourth panel, we see that all methods have problems in identifying small size parameters, however GHR seems to perform lightly better than its competitors in this task. On the other hand, if a blind 95% rule along with the standard t-statistic are used in detecting relevant parameters, GHR leads to over-parametrizing more often than either DR or EML.

5) For VARMA models, given that while EML seems to produce the lowest biases, GHR produces the lowest standard errors, we propose to combine both methods and use, GHR standard errors along with EML point estimates. When this is done, the combined EML procedure is able to detect all relevant parameters in the VARMA(1,1) case. Also, the MSE decreases significatively.

4. CONCLUSIONS

In this paper we generalize the Double Regression estimation method, proposed by Koreisha and Pukkila (1989), for VARMA models. We use a basic idea formulated by Koreisha and Pukkila (1990), that is, the innovations associated to a univariate ARMA(p,q) model will differ from the residuals obtained from a long autoregression. We generalize this idea to the multivariate case, but instead of assuming that residuals and innovations differs each other in a white noise process, we derive the stochastic structure of that difference which turns to be a general VARMA model.

By taking into account the previous result we propose a GLS estimation procedure that we call the Generalized Hannan Rissanen method.

Simulations results indicate that the GHR procedure performs better than the DR method and similar to Mauricio (1995) Exact Maximum Likelihood procedure. It increases the precision of parameters estimates and helps to better identify significant nonzero parameters. This feature is particularly important in the case of low parameters values.

REFERENCES

HANNAN, E.J. and KAVALIERIS, L. (1984), A method for autoregressive-moving average estimation. *Biometrica* 71, 273-80.

HANNAN, E.J. and RISSANEN, J. (1982), Recursive estimation of mixed autoregressive moving average order. *Biometrica* 69, 81-94.

JUDGE, G.G., HILL, R.C., GRIFFITHS, W.E., LÜTKEPOHL, H. and LEE, T. (1982), *Introduction to the Theory and Practice of Econometrics*. New York: John Wiley & Sons.

KOREISHA, S.G. and PUKKILA, T.H. (1989), Fast linear estimation methods for vector autoregressive moving-average models. *Journal of Time Series Analysis*, 10, 325-39.

KOREISHA, S.G. and PUKKILA, T.H. (1990), A generalized least-squares approach for estimation of autoregressive moving-average models. *Journal of Time Series Analysis*, 11, 139-51.

MAURICIO, J.A. (1995), Exact maximum likelihood estimation of stationary vector ARMA models. *Journal of the American Statistical Association*, 90, 282-291.

REINSEL, G.C., BASU, S. and YAP, S.F. (1992), Maximum likelihood estimators in the multivariate autoregressive moving average model from a generalized least squares viewpoint. *Journal of Time Series Analysis*, 13, 133-45.

LÜTKEPOHL, H. (1993), *Introduction to Multiple Time Series Analysis*, Berlin: Springer-Verlag (2nd Ed.)

LÜTKEPOHL, H. and D.S. POSKITT. (1996), Specification of Echelon-Form VARMA models, *Journal of Business & Economic Statistics*, 14, 1, 69-79.

SPLIID, H. (1983), A fast estimation method for the vector autoregressive moving average model with exogenous variables. *Journal of the American Statistical Association*, 78, 843-49.

NOTES

1 Consider

$$\hat{\Sigma}_{GHR} = [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecH_N}(\hat{U}'\otimes I_k)]^{-1} \tag{61}$$

and

$$\hat{\Sigma}_{DR} = [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecA_N}(\hat{U}'\otimes I_k)]^{-1} \tag{62}$$

If

$$[(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecA_N}(\hat{U}'\otimes I_k)]^{-1} - [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecH_N}(\hat{U}\otimes I_k)]^{-1} \tag{63}$$

is positive semidefinite then

$$[(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecH_N}(\hat{U}'\otimes I_k)] - [(\hat{U}\otimes I_k)\hat{\Sigma}^{-1}_{vecA_N}(\hat{U}\otimes I_k)] \tag{64}$$

is positive semidefinite, that is,

$$[(\hat{U}\otimes I_k)(\hat{\Sigma}^{-1}_{vecH_N} - \hat{\Sigma}^{-1}_{vecA_N})(\hat{U}'\otimes I_k)] \tag{65}$$

is positive semidefinite, and

$$\hat{\Sigma}_{vecA_N} - \hat{\Sigma}_{vecH_N} \tag{66}$$

is positive semidefinite. Also, if **(66)** is positive semidefinite so is **(63)**.

2 Koreisha and Pukkila (1989) simulated 50 realizations.

3 We thank J.A. Mauricio for estimating the models in this paper models with his EML algorithm. True parameters values were used to initialize the EML algorithm.

TABLE I

Summary of simulation results VMA(1), K=5, N=100, 100 replications

$$\theta_1 = \begin{bmatrix} 0 & 0 & 0 & 1.1 & 0 \\ 0 & 0 & 0 & 0 & .2 \\ 0 & 0 & 0 & 0 & 0 \\ -.55 & 0 & 0 & .8 & 0 \\ 0 & 0 & 0 & 0 & .6 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & & & & \\ .2 & 1 & & & \\ 0 & 0 & 1 & & \\ 0 & 0 & .7 & 1 & \\ 0 & 0 & 0 & -.4 & 1 \end{bmatrix}$$

	GHR					DR					EML				
	-0.07	-0.02	-0.04	1.13	-0.02	-0.02	-0.01	-0.02	1.09	0.00	0.05	0.00	-0.09	1.24	0.05
	0.02	-0.06	-0.01	0.02	0.21	0.02	-0.05	0.01	-0.01	0.20	0.01	0.00	0.01	-0.01	0.23
θ_1	-0.01	-0.00	-0.05	0.02	0.02	0.00	0.00	-0.05	0.02	0.02	-0.02	0.00	0.00	-0.00	0.01
	-0.55	0.00	0.05	0.75	0.00	-0.53	0.01	0.02	0.78	0.02	-0.59	0.00	-0.02	0.85	0.03
	-0.02	-0.00	-0.05	0.01	0.53	-0.01	-0.00	-0.06	0.02	0.55	0.00	0.00	-0.01	0.00	0.62
Mean values of the estimated standard errors															
	(0.08)	(0.08)	(0.13)	(0.14)	(0.10)	(0.11)	(0.11)	(0.16)	(0.18)	(0.13)	(0.07)	(0.07)	(0.10)	(0.11)	(0.09)
	(0.08)	(0.08)	(0.12)	(0.13)	(0.09)	(0.11)	(0.11)	(0.16)	(0.17)	(0.12)	(0.11)	(0.11)	(0.19)	(0.22)	(0.15)
θ_1	(0.08)	(0.08)	(0.12)	(0.13)	(0.09)	(0.11)	(0.11)	(0.16)	(0.18)	(0.13)	(0.10)	(0.10)	(0.14)	(0.14)	(0.11)
	(0.08)	(0.08)	(0.12)	(0.14)	(0.10)	(0.11)	(0.11)	(0.16)	(0.18)	(0.13)	(0.05)	(0.05)	(0.08)	(0.09)	(0.06)
	(0.08)	(0.08)	(0.12)	(0.13)	(0.09)	(0.11)	(0.11)	(0.16)	(0.18)	(0.13)	(0.10)	(0.08)	(0.12)	(0.13)	(0.11)
MSE (%)															
	1.24	0.78	1.77	2.07	1.06	1.15	1.17	2.60	3.09	1.60	0.75	0.48	1.89	3.14	0.95
	0.63	0.99	1.38	1.68	0.86	1.14	1.34	2.50	3.03	1.56	1.29	1.13	3.55	4.82	2.24
θ_1	0.63	0.62	1.66	1.71	0.90	1.15	1.17	2.76	3.16	1.66	1.06	0.98	1.93	2.09	1.20
	0.66	0.67	1.83	2.13	0.95	1.19	1.18	2.61	3.11	1.65	0.36	0.29	0.63	1.04	0.40
	0.67	0.64	1.64	1.72	1.43	1.18	1.18	3.01	3.20	1.82	0.97	0.70	1.34	1.75	1.16
Frequency of significative nonzero values (%)															
	42	39	36	100	35	21	18	14	100	16	30	16	32	100	23
	37	45	40	37	58	20	15	18	18	43	10	9	8	8	46
θ_1	40	44	36	36	44	10	13	16	12	13	10	8	19	14	9
	100	36	32	93	44	100	16	8	95	16	100	20	20	99	20
	40	34	36	33	92	18	16	16	14	94	6	6	8	18	98

TABLE II

Summary of simulation results VMA(2), K=3, N=100, 100 replications

$$\theta_1 = \begin{bmatrix} .7 & 0 & 0 \\ 0 & 1.25 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \theta_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -.75 & 0 \\ 0 & .3 & .6 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & & \\ -.7 & 1 & \\ .4 & 0 & 1 \end{bmatrix}$$

	GHR			DR			EML		
	0.63	-0.00	-0.01	0.66	-0.02	0.01	0.74	0.01	0.01
θ_1	-0.03	1.11	0.01	-0.01	1.21	-0.00	0.06	1.33	-0.04
	0.03	0.02	-0.09	-0.02	-0.04	-0.02	0.02	-0.01	0.02
	0.00	0.01	-0.02	0.00	0.00	-0.01	0.02	0.01	-0.02
θ_2	0.02	-0.71	-0.00	0.03	-0.69	-0.01	-0.08	-0.84	0.05
	0.03	0.34	0.56	0.02	0.31	0.56	0.02	0.31	0.64
Mean values of the estimated standard errors									
	(0.13)	(0.12)	(0.10)	(0.17)	(0.15)	(0.12)	(0.15)	(0.14)	(0.10)
θ_1	(0.15)	(0.13)	(0.11)	(0.17)	(0.16)	(0.12)	(0.10)	(0.09)	(0.06)
	(0.14)	(0.13)	(0.10)	(0.17)	(0.16)	(0.12)	(0.19)	(0.17)	(0.14)
	(0.13)	(0.12)	(0.10)	(0.17)	(0.15)	(0.12)	(0.15)	(0.13)	(0.10)
θ_2	(0.15)	(0.13)	(0.11)	(0.17)	(0.15)	(0.12)	(0.10)	(0.10)	(0.06)
	(0.14)	(0.13)	(0.10)	(0.17)	(0.16)	(0.12)	(0.22)	(0.19)	(0.14)
MSE (%)									
	2.25	1.49	0.93	3.02	2.44	1.52	2.32	1.84	1.05
θ_1	2.28	3.75	1.13	2.94	2.59	1.52	1.30	1.49	0.53
	1.97	1.61	1.86	2.96	2.58	1.53	3.57	2.77	1.93
	1.77	1.50	0.95	2.87	2.39	1.51	2.19	1.75	1.02
θ_2	2.19	1.93	1.11	3.01	2.78	1.53	1.59	1.79	0.63
	2.01	1.74	1.13	2.93	2.41	1.61	4.81	3.72	2.16
Frequency of significative nonzero values (%)									
	93	33	36	97	6	11	98	10	5
θ_1	31	100	28	12	100	14	30	99	21
	29	27	36	9	9	11	11	11	17
	36	34	33	17	15	15	12	13	17
θ_2	42	92	40	25	91	26	35	99	25
	22	67	99	11	54	99	18	64	97

TABLE III

Summary of simulation results VARMA(1,1), K=3, N=100, 100 replications

$$\Phi_1 = \begin{bmatrix} .7 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & .4 & 0 \end{bmatrix}, \quad \theta_1 = \begin{bmatrix} 0 & -1.1 & 0 \\ 0 & .6 & 0 \\ 0 & 0 & -.5 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & & \\ -.7 & 1 & \\ .4 & 0 & 1 \end{bmatrix}$$

	GHR			DR			EML		
	0.76	0.59	0.13	0.72	0.54	0.17	0.68	-0.10	-0.03
ϕ_1	-0.02	-0.31	-0.06	-0.01	-0.28	-0.09	0.01	0.06	0.01
	0.02	0.49	0.22	0.01	0.48	0.20	0.01	0.34	-0.06
	0.11	-0.51	0.08	0.02	-0.57	0.16	0.04	-1.19	-0.05
θ_1	-0.04	0.31	-0.03	0.02	0.35	-0.09	0.01	0.66	0.03
	0.06	0.13	-0.31	0.02	0.10	-0.29	0.03	-0.03	-0.55
Mean values of the estimated standard errors									
	(0.08)	(0.10)	(0.10)	(0.09)	(0.14)	(0.11)	(0.13)	(0.63)	(0.48)
ϕ_1	(0.08)	(0.10)	(0.10)	(0.09)	(0.13)	(0.11)	(0.07)	(0.36)	(0.29)
	(0.09)	(0.11)	(0.11)	(0.09)	(0.14)	(0.11)	(0.16)	(0.50)	(0.35)
	(0.13)	(0.18)	(0.15)	(0.20)	(0.21)	(0.17)	(0.21)	(0.63)	(0.54)
θ_1	(0.12)	(0.18)	(0.15)	(0.19)	(0.21)	(0.17)	(0.15)	(0.37)	(0.36)
	(0.13)	(0.19)	(0.16)	(0.20)	(0.21)	(0.17)	(0.19)	(0.55)	(0.36)
MSE (%)									
	1.07	36.35	2.72	0.83	31.09	4.01	1.67	40.18	23.11
ϕ_1	0.76	11.16	1.39	0.81	9.88	1.98	0.44	13.60	8.15
	0.76	2.55	5.89	0.84	2.47	5.05	2.45	25.04	12.85
	4.43	38.97	2.99	3.84	32.37	5.28	4.69	40.01	29.69
θ_1	3.29	11.88	2.37	3.82	10.51	3.59	2.12	14.05	12.88
	3.88	5.65	6.07	3.94	5.43	7.04	3.69	30.53	13.07
Frequency of significative nonzero values (%)									
	100	95	35	100	91	34	95 [100]*	9 [40]	7 [44]
ϕ_1	14	60	19	3	54	14	3 [1]	8 [19]	7 [28]
	21	89	56	12	90	46	8 [32]	38 [71]	18 [40]
	20	64	27	13	70	27	15 [19]	95 [100]	8 [30]
θ_1	19	40	21	9	40	10	8 [4]	84 [87]	10 [26]
	11	19	48	6	11	41	15 [11]	8 [24]	76 [79]

(*) Results from combining GHR standard errors and EML point estimates appear in brackets